



**PRODUCTIVITY POLARIZATION ACROSS
REGIONS IN EUROPE**

Roberto Basile

Quaderno n. 31 — Luglio 2007

**QUADERNI DEL DIPARTIMENTO
DI ECONOMIA, FINANZA
E STATISTICA**

**Dipartimento di Economia,
Finanza e Statistica**

Università degli Studi di Perugia
Via A. Pascoli — 06123 Perugia
tel. 075 — 585.5279 — fax 075 — 585.5299
e-mail: diec@unipg.it

Productivity Polarization

Across Regions in Europe

Roberto Basile

Institute for Studies and Economic Analyses

Email: r.basile@isae.it,

Address: ISAE, Piazza dell'Indipendenza, 4, 00185 – Rome,

Tel. +39-06-44482874; Fax +39-06-44482249

Abstract. The regional distribution of labor productivity in Western Europe is characterised by a Core-Periphery spatial pattern: high (low) productivity regions are in a proximate relationship with other high (low) productivity regions. Over the last twenty years, intra-distribution dynamics has generated long-run multiple equilibria with the formation of two clubs of convergence. The observed dynamics can be only marginally explained by nonlinear effects in the accumulation of physical capital. In contrast, the joint effect of spatial dependence and nonlinearities in growth behavior plays a key role in determining multiple equilibria and reinforcing polarization of labor productivity.

Keywords: distribution dynamics, convergence, spatial dependence, Europe.

JEL classification numbers: R11, R12, C14, C21.

Acknowledgments

A previous version of this paper was presented at the AIEL Conference in Naples (University Parthenope), the EU-REAL workshop in Milan (Bocconi University) and at seminars in Perugia and Rome (Luiss University). I wish to thank all the participants at those meetings for helpful comments. The usual disclaimers apply.

1. Introduction

In this work we use a continuous state-space approach to analyze the intra-distribution dynamics (thereafter, *IDD*) (Quah, 1997; Magrini, 2004; Pittau and Zelli, 2006) of regional labour productivity in Western Europe (EU15) over the period 1980-2003. In particular, we apply a robust nonparametric conditional density estimator (Hyndman and Yao, 2002) to describe the law of motion of regional labor productivity and compute the ergodic distribution to identify long-run properties of the observed distribution dynamics (Johnson, 2004). Moreover, using a two-step approach (Lamo, 2000; Bandyopadhyay, 2003; Leonida, 2003), we try to estimate the effect of some economic determinants on the long-run distribution.

The advantages of the *IDD* approach with respect to the growth regression analysis are well known (Quah, 2003, 2006a, b). In particular, it allows to examine directly how the whole productivity distribution changes over time and, thus, it is much more informative than the convergence approach developed within the regression paradigm which gives only information on the dynamics of the average economy. However, some important drawbacks also characterize the literature on *IDD*. First, while the regression approach to economic convergence has been improved in many respects over the last decade¹, most of the studies based on the *IDD* approach scantily take into account the recent developments of the literature on conditional density (Hyndman *et al.*, 1996; Fan

¹ Authors have proposed various estimators and econometric procedures to reduce bias and inefficiency of the regression estimates (for example, IV and GMM estimators), to take account of spatial spill-over effects (spatial econometric techniques), to identify nonlinearities in growth behaviour (nonparametric and semiparametric regression techniques) and to reduce model uncertainty (Bayesian approaches) (for a review of the literature, see Durlauf *et al.* 2005).

et al., 1996; Hall *et al.*, 1999), which have highlighted the bias problems associated with the widely used standard kernel method and have proposed new estimators with better statistical properties. In the present paper, we try to overcome this limit by using log-likelihood conditional density estimators with variable bandwidths (Hyndman and Yao, 2002)².

Second, while testing conditional convergence hypotheses is a very common practice within the growth regression analysis, little effort is usually devoted within the *IDD* approach to investigate the determinants of the long-run (ergodic) distribution. Quah (1997) proposed a “conditioning” scheme which allows to analyze the role of a single variable per time. More recently, some attempts to detect the joint effect of many variables appeared in the literature (Lamo, 2000; Bandyopadhyay, 2003; Leonida, 2003). In particular, a two-step approach is applied, where the first step consists of estimating a growth regression model, while in the second step the residuals from that regression are used to simulate end-period labor productivity levels and, thus, to estimate conditional density functions having filtered out the effect of some growth determinants. However, all these studies use linear regression models to estimate the first step, disregarding the presence of nonlinearities in growth behavior widely highlighted in the growth regression literature (Liu and Stengos, 1999; Banerjee and Duflo, 2003; Masanjala and Papageorgiou, 2004; Basile, 2007b). In the present paper

² Only recently, Johnson (2004) and Fiaschi and Lavezzi (2007) have applied kernel density estimators with adaptive bandwidths to solve in some way the bias problem of the standard kernel density estimator with fixed bandwidth. Basile (2007a) compares different conditional density estimators to analyze the cross-sectional distribution dynamics of regional per-capita incomes in Europe and shows that, while the kernel estimator with fixed bandwidth gives evidence of convergence, a modified estimator with variable bandwidth and mean-bias correction provides evidence of strong persistence and lack of cohesion.

we overcome this limitation by applying modern nonparametric regression techniques (Wood, 2006) in order to remove the effect of growth determinants from conditional density estimations.

We propose the two above mentioned methodological improvements (robust conditional density estimators and semiparametric regressions) in order to answer many interesting questions: Are there convergence tendencies within the group of regional economies included in the sample? If not, does one observe any specific distribution pattern? Do high-productivity regions belong to a club of high-productivity economies, while low-productivity regions languish behind (club convergence hypothesis)? What are the factors that help explain the observed dynamics of the entire distribution? In particular, does capital accumulation explain the difference between growth paths of high- and low-productivity regions? Alternatively, what is the role of sectoral specialization and economic agglomeration? Finally, does spatial dependence matter?

The results of the analysis can be summarized as follows. First, the regional distribution of productivity in Europe is characterised by a clear Core-Periphery spatial pattern which contributes to determine a strong and increasing bimodality in the snapshot univariate density: high (low) productivity regions are in a proximate relationship with other high (low) productivity regions. Second, over the last twenty years, *IDD* has generated long-run multiple equilibria with the formation of two clubs of convergence: regions with low levels of labor productivity at the initial period have hardly managed to get close to the European average productivity in 23 years. The observed dynamics can be only marginally explained by nonlinear effects in the accumulation of physical capital. In contrast, the joint effect of spatial dependence and

nonlinearities in growth behavior plays a key role in determining multiple equilibria and reinforcing polarization of labor productivity.

The layout of the paper is the following. In Section 2, we present a univariate analysis of regional labour productivity in Europe. In section 3, we report the results of the *IDD* analysis. In Section 4, we apply the ‘multivariate’ conditioning scheme and discuss the shape of the ergodic distributions computed after having removed the effects of some growth determinants. Conclusions are reported in Section 5.

2. An Exploratory Spatial Data Analysis

Most of the studies on regional convergence consider the per-capita GDP in order to measure regional unbalances. Some authors (Paci, 1997; Lopez-Bazo *et al.*, 2004) have criticized this choice, observing that GDP is measured at the workplace while population at the residence and, thus, the level of per-capita GDP may lead to great distortions in some regions due to the presence of commuting patterns.³ Based on the same considerations, here we analyse regional convergence of labour productivity, defined as the ratio between GVA (Total Gross Value Added) at constant prices 1995 and total employment for a sample of 190 NUTS-2 European regions over the period 1980-2003. Labour productivity levels are normalized with respect to the EU15 average in order to remove co-movements due to the European wide business cycle and trends in the average values. Data are drawn from Cambridge Econometrics (2006 release).

Over the period 1980-2003 the standard deviation of the relative regional labour productivity decreased by 18% (from 0.323 to 0.264), indicating a slight σ -

³ Clear examples of this are Brussels in Belgium, Hamburg or Bremen in Germany, Inner London in the United Kingdom.

convergence. However, standard deviations might mask some important features of the distribution. In fact, the snapshot densities reported in Figure 1 clearly display a bimodal distribution of labour productivity both in 1980 and in 2003, indicating the existence of two clusters of regions, respectively characterised by low and high levels of labour productivity.⁴ The reduction of the standard deviation is the consequence of the decrease in the mass at the extreme tails of the distribution, while the two peaks become more pronounced in the last year.

Figure 1

Even if it is beyond the scope of this paper to identify the exact composition of the two clusters, it remains important to search for spatial patterns in the distribution of labour productivity. For this purpose, we use some measures of global and local spatial dependence as well as different mapping tools. First, Figure 2 shows a choropleth map of the percentile distribution of regional labour productivity.⁵ This map allows highlighting the existence and persistence of a Core-Periphery pattern in the regional distribution of labour productivity in Europe. A notable exception is Ireland (especially

⁴ Univariate densities have been estimated using the local likelihood density estimator (Loader, 1996). A variable bandwidth, selected by generalised cross validation (GCV), has been used together with a tricube kernel function. In order to allow time comparison, we have used the same span ($\alpha = 0.4$) for both years and evaluated the two univariate densities at the same set of data points. The evidence reported in Figure 1 is in line with that obtained by Fiaschi and Lavezzi (2007) using an adaptive kernel density estimator. Following Fiaschi and Lavezzi (2007), we have also applied a bimodality test based on the bootstrap procedure suggested by Efron and Tibshirani (1993). The p-values of this test are equal to 0.004 for the 1980 and to 0.000 for the last year, indicating the rejection of the unimodality hypothesis.

⁵ In order to overcome (or, at least, to limit) the shortcomings of the crude classification of data points in few (usually 4 or 5) classes, we have imposed 100 breaks points - that is one for each percentile point - thus approximating an un-classed choropleth map (Fotheringham *et al.*, 2000).

the Southern and Eastern region), which registered a strong leap forward of its relative productivity level. Sweden and Finnish regions also appear as high-productivity regions.

Figure 2

Whether high (low) productivity regions are in a proximate relationship with other high (low) productivity territories can be more rigorously assessed by using spatial statistics. We have used distances-based binary matrices to calculate the global G statistic of spatial autocorrelation (Getis and Ord, 1992) defined as

$$G(d) = \frac{\sum_i \sum_j w_{ij}(d) x_i x_j}{\sum_i \sum_j x_i x_j} \quad (1)$$

where x_i (x_j) is the value of labor productivity at regions i (j) and w_{ij} are the elements of the binary spatial weights matrix (that is, ones for all neighbours j within lag distance d of i and zeros for all locations greater than d from i). A high positive value of the standardized G statistic, $Z(G) = \frac{G - E(G)}{V(G)^{1/2}}$, indicates that the spatial pattern is dominated by clusters of high values, while a strong negative $Z(G)$ indicates that the spatial pattern is dominated by clusters of low values.

Standardized G variates were computed for lag distances from 424 km (the minimum distance allowing all regions to have at least one link) up to and including 2024 km at 50 km intervals. Figure 3 shows a non-monotonic relation between distance cut-off and global spatial autocorrelation: $Z(G)$ is always positive but it reaches a maximum when the cut-off distance equals 924 km; above that limit, $Z(G)$ values decrease.

Figure 3

Global G statistic is, however, based on the assumption of stationarity or structural stability over space, which is obviously unrealistic in our context. Spatial association must be detected using local spatial autocorrelation indices which allow for local instabilities in overall spatial association. Local G_i^* indices are defined as follows (Ord and Getis, 1995):

$$G_i^*(d) = \frac{\sum_j w_{ij}(d) x_j}{\sum_j x_j} \quad \forall j \quad (2)$$

with $w_{ii} \neq 0$. In our context, G_i^* is a measure of local clustering of labor productivity around region i . If high (low) values of x tend to be clustered around i , the standardized G_i^* will be positive (negative). No longer committed to the global pattern, local G_i^* statistics are free to characterize the spatial autocorrelation of attribute values located within a distance of each target value. Figure 4 shows standardized G_i^* variates for lag distances of 424 and 924 km for 1980 and 2003. A typical Core-Periphery structure clearly emerges for both years: a cluster of high-productivity regions is located in the Centre of Europe (black color), while groups of peripheral regions are characterized by negative standardized G_i^* scores (grey color). Regions with a white color are those with a non-significant value of G_i^* .⁶ For a cut-off distance of 924 km, the cluster of high-productivity regions is much larger, indicating that the territory becomes more homogenous. Being the G_i^* a spatial smoother (or spatial moving average), isolated high-productivity regions, such as Ireland, do not appear in the ‘black’ cluster.

Figure 4

⁶ Under the null hypothesis, standardized G_i^* statistics is asymptotically normally distributed (Ord and Getis, 1995). P-values have been adjusted using the Bonferroni’s method.

3. Intra-distribution mobility

The univariate analysis carried out so far allowed us to identify some interesting features of regional labour productivity data. Nevertheless, that analysis did not give us any information on the changes of the relative position of various regions in the cross-section distribution of labor productivity over time. However, this issue is relevant for assessing the evolution of regional disparities. In order to address this drawback, it is necessary to examine the intra-distribution mobility during the study period following the transition dynamics approach developed by Quah (1993, 1996a, 1996b, 1996c, 1997, 2006).⁷

Given the distribution of regional productivity at time t and its associated probability measure, ϕ_t , this approach consists of describing the law of motion of the stochastic process $\{\phi_t, t \geq 0\}$. If this process is assumed to be first-order Markov, than the law of motion for $\{\phi_t, t \geq 0\}$ can be modeled as an autoregression:

$$\phi_{t+\tau}(y) = \int_0^{\infty} f_{\tau}(y|x)\phi_t(x)dx \quad (3)$$

where $f_{\tau}(y|x)$ is the expected density of y (the productivity levels at time $t+\tau$) conditional upon x (the productivity levels at time t). In other words, the conditional density $f_{\tau}(y|x)$ describes the probability that a given region moves to a certain state of relative productivity given that it has a certain relative productivity level in the initial

⁷ “Distribution dynamics considers not just the time-path of income distributions – each distribution treated as a point-in-time snapshot – but also a law of motion or a mechanism for how the distribution at one time point evolves into that at a later time” (Quah, 2006, p. 14)

period. For analyzing *IDD*, a researcher must estimate $f_\tau(y|x)$ and visualize the output, that is the shape of the productivity distribution at time $t+\tau$ over the range of productivity levels observed at time t .

If the transition density function is time-invariant, then we can compute the ergodic distribution (Johnson, 2004):

$$\phi_\infty(y) = \int_0^\infty f_\tau(y|x)\phi_\infty(x)dx \quad (4)$$

This function describes the long-term behavior of the productivity distribution: it is the density of what the cross-region productivity distribution tends towards, should the system continue along its historical path (Quah, 2006).

Operationally, the *transition dynamics approach* consists of estimating and visualizing the conditional density of y given x . The most popular approach within the *IDD* literature is the kernel density estimator with fixed bandwidths. However, this estimator has some undesirable bias properties (Hyndman *et al.*, 1996) which, in the context of the *IDD* analysis, might bring to get, for example, evidence of convergence while there is persistence (Basile, 2007a). Fortunately, more robust estimators have recently been developed in the literature. In particular, Hyndman and Yao (2002) have proposed a local linear conditional density estimator which is a conditional version of Loader's (1996) density estimator used in section 2 (see Appendix 1). In the present paper, we use this approach to estimate the conditional density of regional labour productivity at 2003 (the last year) given the distribution at 1980.

The results are plotted in Figure 5. These graphical methods for visualizing conditional density estimates, developed by Hyndman *et al.* (1996) and Hyndman (1996), are not common in the literature of *IDD* and, thus, a preliminary discussion on

their features is necessary.⁸ The first plot, called the “*stacked conditional density plot*” (figures 5A), displays a number of conditional densities plotted side by side in a perspective plot.⁹ It facilitates viewing the changes in the shape of the distributions of the variable observed for the 2003 over the range of the same variable observed for the 1980. In other words, like a row of a transition matrix, each univariate density plot describes transitions over the analyzed period from a given productivity value in 1980. Hyndman *et al.* (1996) note that this plot is “much more informative than the traditional displays of three dimensional functions since it highlights the conditioning” (p.13).

Figure 5A and 5B

The second type of plot proposed by Hyndman *et al.* (1996) is the “*highest conditional density region*” (*HDR*) plot (figures 5B). Each vertical band represents the projection on the xy plan of the conditional density of y on x . In each band the 25% (the darker-shaded region), 50%, 75% and 90% (the lighter-shaded region) *HDRs* are reported. A high density region is the smallest region of the sample space containing a given probability. These regions allow a visual summary of the characteristics of a probability distribution function. In the case of unimodal distributions, the *HDRs* are exactly the usual probabilities around the mean value; however, in the case of multimodal distributions, the *HDR* displays different disjointed subregions.

⁸ All of the studies on *IDD* which make use of nonparametric stochastic kernel density estimators provide three-dimensional perspective plots and/or the corresponding contour plots of the conditional density to describe the law of motion of cross-sectional distributions. In such a way, they treat the conditional density as a bivariate density function, while the latter must be interpreted as a sequence of univariate densities of relative productivity levels conditional on certain initial levels.

⁹ Since the conditional density plot has been evaluated on an equispaced grid of 100 values over the range of x and y directions, figure 5A displays 100 stacked univariate densities.

The *HDR* plot is particularly suitable to analyze *IDD*. If the 45-degree diagonal crosses the 25-50% *HDRs*, it means that most of the elements in the distribution remain where they started (there is '*persistence*'). If the horizontal line traced at the one-value of the vertical axis crosses *all* the 25-50% *HDRs*, we can say that there is '*global convergence*' towards equality. If the vertical line traced at the one-value of the horizontal axis crosses *all* the 25-50% *HDRs*, we can say that there is '*global divergence*'. Finally, the presence of nonlinearities in the modal regression functions (shown in the plot as bullets) can be interpreted as an evidence in favor of the '*convergence club*' hypothesis, according to which regions catch up with one another but only within particular sub-groups.

Figure 5 shows the existence of two convergence clubs: regions sufficiently close to each other converge towards each other. The first club is composed of 50 regions with a relative productivity level at 1980 ranging between 0.26 and 0.70 times the EU15 average; the second club is composed of 54 regions with a relative productivity level at 1980 ranging between 1.15 and 1.77 times the EU15 average. At the upper tails of the distribution, we can observe some bimodality in the conditional density function suggesting the existence of some dualistic behaviour: a few regions with very high productivity levels do not converge to any level; rather they tend to persist in their relative positions. Finally, an intermediate area, composed of 80 regions with a relative productivity level at 1980 ranging between 0.70 and 1.15 times the EU15 average, can be classified as an area of persistence. Finally, six regions can be classified as outliers. In particular, Ireland - which started from a relative value of 0.70 and ended up with a relative value of 1.25 - seems to belong neither to Club 1 nor to the Area of persistence. Figure 6 clearly shows that the two clubs have also a core-periphery spatial pattern.

Figure 6

The shape of the ergodic distribution (Figure 1) suggests that, in the long run the European system might tend towards some reduction of regional unbalances even if the twin-peaks property remains: the first peak of the stationary distribution occurs at a slightly higher relative productivity level than that of the initial distribution; the second peak is much higher than that of both the initial and the final distributions and a decrease in the mass at the extreme tails of distribution occurs.¹⁰

A relevant question is how fast the actual distribution approaches the steady state one. This can be assessed from the system's *half life* obtained as $-\log 2 / \log(\text{eigen_2})$, where *eigen_2* is the second eigenvalue of the one-step probability matrix. In our case, the half life is 60 years indicating a rather low pace.

4. The growth determinants

4.1 The conditioning scheme

The analysis carried out so far can be interpreted as a test of the hypothesis of “*absolute convergence*”, since it does not control for the heterogeneity in the structural characteristics of the regions (in terms of technologies, saving rates, sectoral specialization, spatial dependence and so on; see Galor, 1996). Having rejected such hypothesis and having assessed that European regions tend towards different long-run equilibria (the “*club convergence*” hypothesis cannot be rejected), it remains to test the

¹⁰ The ergodic distributions have been computed starting from the transition matrices extracted from each conditional density estimation. In order to compare univariate density functions (at 1980 and 2003) and the ergodic distribution, conditional densities have been evaluated at the same data points at which the initial density function was firstly evaluated.

“*conditional convergence*” and the “*club conditional convergence*” hypotheses, that is we have to understand why low-productivity regions do not tend to converge (in the long run) with high-productivity regions. In other words, our task is now to identify those factors that determine the formation of club convergence. Removing the effect of these factors, the evidence of bimodality in the ergodic distribution should disappear.

Some recent studies on *IDD* have already proposed interesting methodologies to remove the effect of some determinants of economic growth from the realized mobility dynamics across a sample of economies (Lamo, 2000; Leonida, 2003; Cheshire and Magrini, 2006). All these studies have used a two-step procedure consisting of, first, estimating a linear parametric growth regression and, then, using the residuals from this regression to simulate end-period log-relative labor productivities which, through the estimation of the conditional density function, enable to analyze the effect of different variables in shaping the dynamics of cross-regional distribution of labor productivity.

Let us describe this procedure more formally. First, define with $\ln y$ and $\ln x$ the natural logarithmic transformation of labour productivity levels at the final and the initial years, respectively. Thus, the growth rate of labour productivity can be expressed as $\gamma = (\ln y - \ln x) / \tau$. Now, note that the conditional density function, $f(y|x)$, can be written as $f(\exp(\ln x + \tau\gamma) | x)$. We can use this formulation to study the effect of any explanatory variable on the *IDD* of labour productivity, by defining

$$f(\exp(\ln x + \tau(\gamma - \hat{\gamma})) | x) \quad (5)$$

where $\hat{\gamma}$ is the growth rate predicted from a regression model.

This approach is much more appealing than the original conditioning scheme proposed by Quah (1997), since it allows to conditioning out the effect of many

variables jointly. A shortcoming of this method is, however, evident: it imposes linearity in the functional form of the growth regression equation within a (flexible) nonparametric framework aimed (among other things) at identifying nonlinearities and convergence clubs. In the present paper, we propose to use, for the first step, a nonparametric additive model which allows identifying nonlinearities in growth behavior.

4.2 The specification of the growth regression models

The choice of variables in the growth regression model varies greatly in the empirical literature, therefore giving rise to the ‘model uncertainty’ problem. A recent and very promising line of research for identifying effective regressors is based on Bayesian Model Averaging (Fernández *et al.*, 2002). This issue goes beyond the scope of the present paper. However, we have tried to reduce potential misspecification error from choice of explanatory variables, by incorporating variables whose explanatory power was established to be robust in other works on European regions convergence. In particular, we added to the standard Solow-type conditioning factors (physical capital accumulation, employment growth and initial conditions), measures of sectoral specialization (Paci and Pigliaru, 1999), employment agglomeration (Ciccone and Hall, 1996; Fiaschi and Lavezzi, 2007) and spatial dependence (Lopez-Bazo *et al.* 2004; Basile, 2007b).¹¹ Therefore, we consider five different nested and non-nested models:

¹¹ Spatial dependence and sectoral specialization may be also important sources of multiple steady-state equilibria and club convergence. For instance, in a two-sector overlapping-generation model in which a distinction is made between consumption goods and investment goods (Galor, 1996), multiplicity of steady-state equilibria occurs under a less restrictive set of assumptions than those required in the one-sector model. Ertur and Koch (2006) propose an augmented Solow-type model with spatial externalities (spatial knowledge spillovers) between economies and provide an equation for the steady state level as

$$\text{Eq. 1} \quad \gamma = \alpha_1 + \beta \ln x + \phi \ln \left[\frac{i_k}{n+g+\delta} \right] + \varepsilon_1 \quad (6)$$

$$\text{Eq. 2} \quad \gamma = \alpha_2 + s_1 (\ln x) + s_2 \left(\ln \left[\frac{i_k}{n+g+\delta} \right] \right) + \varepsilon_2 \quad (7)$$

$$\text{Eq. 3} \quad \gamma = \alpha_3 + s_3 (\ln x, W \ln x) + s_4 \left(\ln \left[\frac{i_k}{n+g+\delta} \right], W \ln \left[\frac{i_k}{n+g+\delta} \right] \right) + s_5 (W \gamma) + \varepsilon_3 \quad (8)$$

$$\text{Eq. 4} \quad \gamma = \alpha_4 + s_6 (\ln x) + s_7 \left(\ln \left[\frac{i_k}{n+g+\delta} \right] \right) + s_8 (\ln agri) + s_9 (\ln dens) + \varepsilon_4 \quad (10)$$

$$\begin{aligned} \text{Eq. 5} \quad \gamma = & \alpha_5 + s_{10} (\ln x, W \ln x) + s_{11} \left(\ln \left[\frac{i_k}{n+g+\delta} \right], W \ln \left[\frac{i_k}{n+g+\delta} \right] \right) \\ & + s_{12} (\ln agri, W \ln agri) + s_{13} (\ln dens, W \ln dens) + s_{14} (W \gamma) + \varepsilon_5 \end{aligned} \quad (9)$$

The former (Eq. 1) is the standard “*linear Solow model*”. The first term on the right hand side, $\ln x$ (the per worker GDP in 1980), captures the effect of interregional differences in initial aggregate productivity on interregional differences in growth rates.

The second term, $\ln \left[\frac{i_k}{n+g+\delta} \right]$, is the effective net physical capital accumulation rate.

It captures the combined effect of the investment ratio (i_k , investment to GDP over the period 1980-2003), the average grow rate of employment from 1980 to 2003 (n), the depreciation rate (δ) and the growth rate of technology (g). Following many other growth regression studies, we assume that $(g + \delta) = 0.05$.¹²

well as a conditional convergence equation characterized by parameter heterogeneity: since knowledge spill-overs are local - rather than global - in scope, multiple equilibria (and, thus, convergence clubs) do emerge.

¹² Eq. 1, as well as all the other model specifications, imposes a restriction in the coefficients of $\ln \left[\frac{i_k}{n+g+\delta} \right]$ and $\ln [n+g+\delta]$. This restriction has been formally tested in the case of Eq. 1 through a standard F-

The second specification (Eq. 2) is a “*nonlinear Solow model*”, where both $\ln x$ and $\ln \left[\frac{i_k}{n + g + \delta} \right]$ enter as smooth additive terms. Assuming parameter homogeneity in growth regression is equivalent to assuming that the effect of a change in a particular variable (such as the saving rate) on economic growth is the same across regions. This concern has been addressed by estimating more general models that allow for multiple regimes, parameter heterogeneity and nonlinearities (Durlauf and Johnson, 1995; Liu and Stengos, 1999; Durlauf *et al.*, 2001; Kurtellos, 2001; Masanjala and Papageorgiou, 2004; Basile, 2007b). In a nutshell, nonlinearities in growth regressions have at least three possible interpretations: a) existence of multiple steady-states, b) omitted growth determinants and c) nonlinearity in the production function. In particular, Azariadis and Drazen (1990) consider the consequences of a discontinuity in the aggregate production function for aggregate economies. This discontinuity means that the steady-state behaviour of a given economy depends on whether its initial capital stock is above or below a certain threshold.

The third specification (Eq. 3) is the “*spatial nonlinear Solow model*”. It is theoretically based on recent contributions to the growth literature (Lopez-Bazo *et al.*, 2004; Ertur and Koch, 2006), which have proposed modified versions of the neoclassical growth model that include spatial externalities between economies in order to take account of the neighbourhood effects (that is ‘spatial diffusion with frictions’) on growth and convergence processes. In particular, Ertur and Koch (2006) propose a test. The null hypothesis of statistically equal coefficients cannot be rejected. The F statistic is equal to 0.491 with a p-value of 0.484. An approximate F-test has also been carried out to test the same hypothesis for each of the nonparametric additive models (Eq.2-5). The results always confirm that the parameter restriction cannot be rejected.

growth regression specification with parameter heterogeneity. Following this literature, Basile (2007b) has already employed a semiparametric unrestricted Spatial Durbin model to analyse the growth behaviour of 155 European regions in the period 1988-2000. All these studies suggest that the growth rate of a region can be a (nonlinear) function of its initial conditions ($\ln x$) and of the initial conditions of its neighbours ($W \ln x$).¹³ It is also a (nonlinear) function of reproducible factors accumulation rates

observed within the region $\left(\ln \left[\frac{i_k}{n + g + \delta} \right] \right)$ and in its neighbours $\left(W \ln \left[\frac{i_k}{n + g + \delta} \right] \right)$. As

in Basile (2007b), the effect of these variables are captured by introducing nonparametric interaction terms in the model. Eq. 3 includes also the spatial lag of the dependent variable, $(W\gamma_y)$, that is the weighted average of the growth rates of neighbouring regions. As well known, this term is correlated with the disturbance term, ε_3 , and both maximum likelihood (Anselin, 1988) and instrumental variable (IV) (Kelejian and Prucha, 1998) estimators have been suggested in the literature as consistent procedures for estimating this model. As shown below, here we apply a two-stage approach in order to control for the endogeneity of this term, using higher-order

¹³ W is a standardized spatial weights matrix. The elements w_{ii} on the main diagonal set to zero, while the elements w_{ij} indicate the way region i is spatially connected to the region j . In particular, $w_{ij} = d_{ij}^{-2}$ if $d_{ij} < \bar{d}$ and $w_{ij} = 0$ if $d_{ij} > \bar{d}$, where d_{ij} is the great circle distance between the centroids of region i and region j and \bar{d} is the cut-off (924 kms) corresponding to the highest value of spatial autocorrelation (see Figure 3). The spatial weights matrix is row-standardized such that the elements in each row sum up to one.

spatial lags as exogenous instruments.¹⁴ The presence of this endogenous term has also important implications for the interpretation of the interaction effects $s_3(\ln x, W \ln x)$ and $s_4\left(\ln\left[\frac{i_k}{n+g+\delta}\right], W \ln\left[\frac{i_k}{n+g+\delta}\right]\right)$. In presence of a significant and positive effect of $(W\gamma_y)$, indeed, a spatial multiplier enters the reduced form of the model and spillovers become global, reflecting not only effects from direct neighbors, but also from the neighbors' neighbors.¹⁵

Eq. 4 includes some other terms in the nonlinear Solow model, namely the average share of agriculture value-added on total value-added over the period 1980-2003, $s_8(\ln agri)$, and the average share of total employment on the area surface of the region (in squared km) over the period 1980-2003, $s_9(\ln dens)$. Finally, Eq. 5 includes all variables in the model, so that all other specifications are nested on it.

Additive models specified in Eq. 2-5 are estimated using bivariate thin-plate regression splines and applying the method described in Wood (2006) that allows integrated smoothing parameter selection via *GCV* (see Appendix 2). This method (implemented in the R package *mgcv*) helps overcome the difficulties of model selection

¹⁴ The endogeneity problem in our cross-section analysis may be present also because other variables used as regressors (such as physical capital accumulation, employment density and agriculture specialization) may be influenced by the same factors that affect output. We might use an IV approach to control also for these endogeneity sources. However, in these cases, treatment of endogeneity problems is more difficult because of the absence of internal instruments and, as observed by Brock and Durlauf (2000), studies using IV to address endogeneity are not convincing as their choice of instruments does not meet the necessary exogeneity requirements.

¹⁵ See Anselin (2004) for the correct interpretation of spatial models and Basile (2007b) for the interpretation of semiparametric unrestricted Spatial Durbin models.

typical of the AM framework based on back-fitting developed by Hastie and Tibshirani (1990).¹⁶

4.3 Regression results

Tables 1 and 2 report the results and a series of diagnostic tests for the different models. The proportion of deviance explained ranges from 47.7% (linear Solow model – not reported) to 83.3% (model 5), while the *GCV* score reaches the lowest level with model 5, clearly suggesting that the most general specification encompasses all the others. Moreover, the hypotheses of normality, constant variance and no spatial dependence in the residuals cannot be rejected only in the case of model 5. All this suggests that spatial dependence and nonlinearities must be explicitly taken into account in order to avoid misspecification problems and that sectoral specialization and employment density partially contribute to explain heterogeneity in regional growth behavior in Europe.

Tables 1 and 2

The linear coefficients of model 1 (the standard linear Solow growth model) have the expected signs and are significantly different from zero. The coefficient β associated to $\ln x$ is -1.092 , while the coefficient ϕ associated to $\ln \left[\frac{i_k}{n + g + \delta} \right]$ is 0.101 . The *F*-tests for the overall significance of the smoothed terms in AMs 2-5 have p-values lower than 0.05, while the number of effective degrees of freedom (e.d.f.) suggests that the

¹⁶ Apart from the semiparametric approach used here as well as in some other growth analyses (such as Liu and Stengos, 1999), at least other three methods have been used in the growth regression literature to search for parameter heterogeneity: the regression trees approach (Durlauf and Johnson, 1995), the threshold estimator (Masanjala and Papageorgiou, 2004) and the varying coefficient model (Kourtellis, 2001).

relationship between growth rates and Solow-growth determinants is far from being linear, except for the terms $\hat{\beta}_6(\ln x)$ estimated with only 1 e.d.f.

To save space, we only discuss graphical results for the full model 5 (Figure 7). The vertical axis reports the scale of productivity growth rates; the axes on the plane report the scale of each independent variable and of its correspondent spatial lag. All spatial lags are computed using a row-standardized spatial weights matrix, whose elements w_{ii} on the main diagonal are set to zero whereas elements $w_{ij} = d_{ij}^{-2}$ if $d_{ij} < \bar{d}$ and $w_{ij} = 0$ if $d_{ij} > \bar{d}$, with d_{ij} the great circle distance between the centroids of region i and region j and \bar{d} the cut-off (924 kms) corresponding to the highest value of spatial autocorrelation (see Figure 3).

Figure 7

Figure 7A shows the estimated effect of the interaction between $\ln x$ and its spatial lag on the growth rate of labour productivity. It clearly suggests that regions surrounded by higher productivity regions have higher expected growth rates than regions with a lower-productivity neighbourhood. Thus, while very low-productivity regions have generally higher expected growth rates, as it is predicted by the Solow growth model, those with high-productivity neighbours have the highest rates of growth. Moreover, even very high-productivity regions (which are closer to their steady state and, thus, have lower margins for catching up) have chance to grow faster when surrounded by high-productivity regions. Finally, as expected, the effect of the interaction between $\ln x$ and $W \ln x$ is also characterized by some nonlinearities. In particular, the marginal effect of $\ln x$ allows to identify two negatively-sloped segments, indicating two groups of regions converging towards different steady states, and a zero-sloped segment in the

middle, indicating the presence of a non-converging group of middle productivity regions.

Figure 7B shows the marginal effect of the interaction between $\ln\left[\frac{i_k}{n+g+\delta}\right]$ and $W\ln\left[\frac{i_k}{n+g+\delta}\right]$. First, some nonlinearities in the effect of the rate of capital accumulation are clearly detected: an increase in the rate of capital accumulation is associated with an increase in growth rate only when $\ln\left[\frac{i_k}{n+g+\delta}\right]$ is above a certain threshold, as suggested by Azariadis and Drazen (1990). Moreover, the growth rate of a region is also a positive function of the capital accumulation rate in the neighbours, as suggested by Ertur and Koch (2006).

Figures 7C and 7D show the marginal effects of the average share of agriculture value-added on total value-added and of the employment density, respectively. As expected, both terms have a significant effect on growth. In particular, agriculture specialization has a linear negative effect, while the interaction between economic agglomeration and its spatial lag is positive and characterized by strong nonlinearities.

Model 5 includes another term, $s_{14}(W\gamma)$, which measures the smooth spatial externalities effect. As we already said, this term cannot be considered as exogenous. Thus, miming the spatial two-stage least-squares procedure (Kelejian and Prucha, 1998), we have adopted an instrumental variable approach, using spline transformations of spatially lagged exogenous variables as instruments for the spatially lagged

dependent variable.¹⁷ The F test suggests that this term has a significant nonlinear effect on the growth rate and, as expected, the spatial externality curve turns to be positively sloped (Figure 6E). As mentioned above, a positive effect of $W\gamma$ implies that a change in an exogenous variable - such as, for example, $\left(\ln \left[\frac{i_k}{n+g+\delta} \right], W \ln \left[\frac{i_k}{n+g+\delta} \right] \right)$ - as well a random shock in a specific region affect not only the growth performance of that region (local spill-over), but also the growth performance of all other regions in the EU system through a spatial multiplier mechanism (global spill-over). However, it is worth mentioning that, for the estimation of this term, we have imposed some restrictions on the parameters in order to allow local elasticities to be lower than one along the overall range of $W\gamma$ (in fact, the slope of the curve is lower than that of the 45-degree diagonal), in line with the literature on spatial autoregressive models (Anselin, 1988). These restrictions are necessary to avoid explosive spatial multiplier effects.

4.4 Conditioned ergodic distributions

The econometric results discussed above provide some evidence in favour to the hypothesis of nonlinearities in the effect of initial conditions and capital accumulation, thus suggesting that the linear Solow growth model suffers from misspecification problems. They have also highlighted the importance of sector specialization and of

¹⁷ Specifically, we have used smooth terms of the second-order spatial lags of $\ln x$, $\ln \left[\frac{i_k}{n+g+\delta} \right]$,

$\ln(agri)$ and $\ln(dens)$. F-tests in IV specifications (model 3 and 5) always indicate that instruments are jointly significant in the first stage (see Table 2). Moreover, overidentification tests using the Sargan statistic cannot reject the null hypothesis of zero correlation of instruments with second-stage residuals at the 5% level.

economic agglomeration. Moreover, some specifications have allowed identifying the effect of spatial dependence, confirming the prediction of recently developed spatial growth models (Ertur and Koch, 2006).

This section reports the results of ergodic distributions computed after having removed the effect of growth determinants. In practice, we have firstly re-estimated Eq. 2-5 without the smooth term $s(\ln x)$ and the intercept in order to compute the prediction \hat{y} and, thus, estimate conditional densities as in (5). Then, we have estimated ‘conditioned’ ergodic distributions using the transition matrices extracted from each conditional density estimation and compared them with the ‘unconditioned’ ergodic distribution (Figure 8).

Figure 8

Figure 8A reports the ergodic distribution computed after having removed the smooth effect of capital accumulation from the actual productivity growth rate, while Figure 8C displays the ergodic distribution conditioned also to sectoral specialization and economic agglomeration. The two graphs suggest that all these factors only partially explain the long-run distribution of regional labor productivity: the shape of the ergodic distribution remains bimodal, even if the two peaks are less pronounced than the ones displayed in Figure 1.¹⁸

A clear transformation of the ergodic distribution appears only after having filtered out the effect of spatial dependence (Figures 8B and 8D): most of the initially low-productive regions make a transition out of the low-productivity trap and converge to

¹⁸ Bimodality has also been formally tested using the bootstrap procedure suggested by Efron and Tibshirani (1993). The p-values of this test are equal to 0.000 for both the conditioned ergodic distributions displayed in Figures 7A and 7C, indicating the rejection of the unimodality hypothesis.

the neighborhood of the higher attractor.¹⁹ However, residual bumps still characterize the long-run distribution, suggesting that a small fraction of sample regions still remain in a vicious cycle of “low-productivity trap”.

Now, one can wonder whether this result is robust to different choices of the spatial weights matrix and whether the model specification should include some other relevant variables in order to remove residual bumps. An answer to the first question is given by the results of an analysis based on the choice of a spatial weights matrix with a cut-off distance equal to 424 kms (Figure 8E). In this case the ergodic distribution is unimodal, clustered around the single high level attractor although residuals bumps are still evident, confirming that a few regions are caught in a low-productivity trap. So, we can conclude that our results are robust to the choice of the spatial weights matrix. Further effort could be devoted toward finding the factors that remove the evidence of residual bumps. A possibility is to include country factors. Indeed, there is a considerable amount of empirical findings on the role of national factors for the development of regions in the EU (for example, Armstrong, 1995; Rodríguez-Pose, 1999). These studies report evidence of convergence among countries and divergence between regions within countries. Thus, we have estimated again model 5 after using deviations from the country mean. The results, however, are not helpful for our purpose, since the evidence of bimodality is magnified instead of disappearing. Other relevant variables, such as for example public infrastructural investments and human capital investments, are not available for the sample period considered.

¹⁹ The hypothesis of unimodality can now be rejected using the bootstrap procedure: the p-value is equal to 0.12 in both cases.

A further possible robustness check can be carried out by isolating the effect of spatial dependence, that is by estimating model 5 under the restrictive assumption of global linearity in the effect of all growth determinants. In this case the implied ergodic distribution is again strongly bimodal.

5. Discussion and Conclusions

In this paper we have used a continuous state-space approach to analyze the distribution dynamics of regional labor productivity in Western Europe over the period 1980-2003. The evidence on the shape of the ergodic distribution suggests the existence of multiple equilibria in regional growth behavior in Europe with the formation of two clubs of convergence, which have also a clear spatial pattern: high productivity regions, mainly located in the core of Europe, tend to converge towards high productivity levels, while most of the peripheral regions seem to be caught in a vicious cycle of “low-productivity trap”.

We have tried to investigate some potential determinants of the low-productivity trap. If the ergodic distribution was determined only by nonlinearities in the effect of net capital accumulation rate (thresholds effects) as suggested by Azariadis and Drazen (1990), we should observe a vanishing of bimodality after having removed the effect of these nonlinear effects. However, the results of our analysis does not corroborate this hypothesis. Using a two-step approach, we have shown that the ergodic distribution still appears bimodal after having filtered out the (nonlinear) effect of net capital accumulation rates, even if the two peaks are less pronounced than in the unconditioned case.

Our findings also suggest that sectoral specialization and economic agglomeration do not fully accomplish to remove the evidence of bimodality, even if these factors have

a significant impact on the ‘average’ growth rate. In contrast, removing also the effect of spatial dependence, a large fraction of regions make a transition out of the low-productivity trap, while only a few economies still remain in the neighborhood of the lower attractor. Thus, we conclude that spatial dependence is primarily responsible for the bimodality in the long-run distribution of labor productivity.

The results of our analysis may contribute to the wide debate on poverty traps (or low-productivity traps) recently reviewed, for example, by Azariadis and Stachurski (2005) and by Bowles, Durlauf and Hoff (2004). This literature has identified three broad kinds of explanations for the poverty traps. First, there may be critical thresholds in physical or human capital that must be reached before the forces of traditional competitive theory take hold (Azariadis and Drazen, 1990). Neighborhood effects are a second category of explanation for the persistence of poverty (Sampson and Morenoff, 2004). Third, bad institutions, such as those that make property rights insecure, may entrap entire economies in poverty or low-productivity levels. In the present paper, we have focused on the first two hypotheses. The empirical findings suggest that neither threshold effects in physical capital accumulation nor spatial dependence can be considered as unique responsible for the existence of convergence clubs in the case of European regions. Rather, it is the joint effect of spatial dependence and nonlinearities in growth behavior that play a key role in determining multiple equilibria and reinforcing polarization of labor productivity.

Appendix 1: Local linear conditional density estimators

The most common estimator of the conditional density widely used in the literature of *IDD* is the kernel estimator. Recently, Hyndman *et al.* (1996) have explored its properties. They define:

$$\hat{f}_\tau(y|x) = \frac{1}{b} \sum_{i=1}^n w_i(x) K\left(\frac{\|y - Y_i\|_y}{b}\right)$$

where

$$w_i(x) = K\left(\frac{\|x - X_i\|_x}{a}\right) / \sum_{j=1}^n K\left(\frac{\|x - X_j\|_x}{a}\right)$$

Thus, the conditional density estimator can be interpreted as the Nadaraya-Watson kernel regression of $K\left(\frac{\|y - Y_i\|_y}{b}\right)$ on X_i . As it is well known, the Nadaraya-Watson estimator can have a large bias both on the boundaries of the predictor space, due to the asymmetry of the kernel neighbourhood, and in its interior, if the true function has substantial curvature or if the design points are very irregularly spaced.

Given the undesirable bias properties of the kernel smoother, Hyndman and Yao (2002) have proposed a local likelihood conditional density estimator, which is a conditional version of Loader's (1996) density estimator. Let

$$R(\beta_0, \beta_1; x, y) = \sum_{i=1}^n \left\{ K\left(\frac{\|y - Y_i\|_y}{b}\right) - \exp(\beta_0 - \beta_1(X_i - x)) \right\}^2 K\left(\frac{\|x - X_i\|_x}{a}\right)$$

where $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ is that value of β which minimizes $R(\beta_0, \beta_1; x, y)$. The local linear density estimator at a focal point x is then defined as $\hat{f}(y|x) = \hat{\beta}_0$. This estimator has a smaller bias than the Nadaraya-Watson estimator. All conditional densities in the

present paper have been estimated using Hyndman and Yao (2002). Optimal bandwidths, a and b , have been selected using the method developed by Bashtannyk and Hyndman (2001) based on GCV.

Appendix 2: Additive models with integrated model selection

Additive models (AM) provide a framework for nonparametric and semiparametric modeling. In general the model has a structure something like:

$$y_i = \mathbf{X}_i^* \boldsymbol{\theta} + s_1(x_{1i}) + s_2(x_{2i}) + s_3(x_{3i}, x_{4i}) + \dots + \varepsilon_i \quad \varepsilon_i \sim i.i.d. N(0, \sigma^2) \quad (\text{A1})$$

where y_i is an univariate response continuous variable, \mathbf{X}_i^* is a vector of strictly parametric components, $\boldsymbol{\theta}$ is the corresponding parameter vector and $s_j(\cdot)$ are smooth functions of the covariates, x_j . The estimated function $\hat{s}(\cdot)$ can reveal possible nonlinearities in the effect of x_j .

The most popular approach for estimating AM is the back-fitting algorithm proposed by Hastie and Tibshirani (1990). This approach, however, presents some shortcomings with respect to the issues of model selection and inference. Wood (2000, 2006) and Wood and Augustin (2002) have recently proposed a new method to estimate AM with spline based penalized regression smoothers which allows for automatic and integrated smoothing parameters selection via Generalized Cross Validation (GCV). Wood has implemented this approach in the R package *mgcv*.

In the case of a model containing one smooth function of one covariate ($y_i = s(x_i) + \varepsilon_i$), the penalized regression spline arises as the solution to the following optimization problem:

$$\min \|\mathbf{X}\boldsymbol{\beta} - \mathbf{y}\|^2 + \lambda \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta} \quad (\text{A2})$$

w.r.t. $\boldsymbol{\beta}$ (the parameter vector). $\|\cdot\|$ is the Euclidean norm and \mathbf{S} is a positive semi-definite matrix depending on the basis functions evaluated at x . Given λ (a constant smoothing parameter), the solution to (A2) is:

$$\hat{\boldsymbol{\beta}} = [\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S}]^{-1} \mathbf{X}^T \mathbf{y} \quad (\text{A3})$$

A crucial issue in the use of smoothing splines is the selection of parameter λ , controlling the trade-off between fidelity to the data and smoothness of the fitted spline. GCV is the most common method used to choose the smoothing parameter:

$$GCV(\lambda) = \frac{n \|\mathbf{y} - \mathbf{A}\mathbf{y}\|^2}{[n - \text{tr}(\mathbf{A})]^2} \quad (\text{A4})$$

where \mathbf{A} is the hat matrix for the model being fitted: $\mathbf{X} \left(\mathbf{X}^T \mathbf{X} + \sum_i \lambda \mathbf{S}_i \right)^{-1} \mathbf{X}^T$, and the term $\text{tr}(\mathbf{A})$ gives the estimated degrees of freedom of the model. The best λ is the one that minimizes GCV.

When there are two or more smoothed terms (e.g. $y_i = s_1(x_{1i}) + s_2(x_{2i}) + \varepsilon_i$), the selection of the smoothing parameters becomes less straightforward. Consider first the back-fitting algorithm proposed by Hastie and Tibshirani (1990). It consists of estimating each term by iteratively smoothing partial residuals from the AM w.r.t. the covariate that the smooth relates to. Thus, given the bandwidth of the smoothers, the estimation of smooth terms becomes straightforward with back-fitting. However, “estimation of that bandwidth is hard to integrate into a back-fitting approach” and the choice of the degree of smoothness of each term in the model becomes arbitrary (Wood and Augustin, 2002, p. 2). To overcome this problem, Wood (2000) provides a methodology to choose automatically multiple smoothing parameters by GCV, as in the

single-penalty case. First, he suggests to write the model fitting problem with an extra “overall” smoothing parameter (ρ) controlling the trade-off between model fit and overall smoothness:

$$\min \left\| (\mathbf{X}\boldsymbol{\beta} - \mathbf{y}) \right\|^2 \rho + \sum_{j=1}^m \lambda_j \boldsymbol{\beta}^T \mathbf{S}_j \boldsymbol{\beta} \quad (\text{A5})$$

w.r.t. $\boldsymbol{\beta}$ subject to the linear constraint $\mathbf{C}\boldsymbol{\beta} = 0$, where \mathbf{C} is a matrix of known coefficients defining the constraints. The smoothing parameters, ρ and $\boldsymbol{\lambda}$, are estimated by minimizing the GCV score:

$$GCV(\rho, \boldsymbol{\lambda}) = \frac{n \left\| (\mathbf{y} - \mathbf{A}(\rho, \boldsymbol{\lambda})\mathbf{y}) \right\|^2}{\left[n - \text{tr}(\mathbf{A}(\rho, \boldsymbol{\lambda})) \right]^2} \quad (\text{A6})$$

w.r.t. the relative smoothing parameters, $\boldsymbol{\lambda}/\rho$. Problems (A4) and (A6) are solved iteratively:

1. given the current estimates of the relative smoothing parameters, estimate the overall smoothing parameter using single smoothing parameter methods;
2. given the overall smoothing parameter, update the logarithms of the relative smoothing parameters simultaneously using Newton’s method.

Therefore, with this method, the smoothing parameters for each smooth term in the model are chosen simultaneously and automatically as part of model fitting by minimizing the GCV score of the whole model.

So far, the approach for estimating an AM with the automatic model selection developed by Wood (2000) has been described for the simple case of one dimensional basis functions. Wood and Augustin (2002) and Wood (2003) have extended this approach to the cases of multi-dimensional bases, in particular to thin-plate regression splines and to tensor products. Specifically, Wood (2006) recommends to use thin-plate regression splines

for smooth interactions of quantities measured in the same units, while he suggests to use tensor products for smooth interactions of quantities measured in different units, or when very different degrees of smoothing are appropriate relative to different covariates.

Table 1 – Results of additive models

	Soothed terms	F test (p-values)	e.d.f.	Deviance	AIC	GCV × 1000
Model 2	$s_1(\ln x)$	20.6 (0.000)	3.2	59.8	-1,522	0.019
	$s_2\left(\ln\left[\frac{i_k}{n+g+\delta}\right]\right)$	2.6 (0.009)	3.8			
Model 3	$s_3(\ln x, W \ln x)$	4.3 (0.000)	10.3	68.1	-1,548	0.017
	$s_4\left(\ln\left[\frac{i_k}{n+g+\delta}\right], W \ln\left[\frac{i_k}{n+g+\delta}\right]\right)$	1.5 (0.043)	4.3			
	$s_5(W\gamma)$	10.8 (0.000)	1.3			
Model 4	$s_6(\ln x)$	164.5 (0.000)	1.0	68.2	-1,549	0.017
	$s_7\left(\ln\left[\frac{i_k}{n+g+\delta}\right]\right)$	3.8 (0.000)	3.5			
	$s_8(\ln agri)$	3.5 (0.000)	7.0			
	$s_9(\ln dens)$	2.7 (0.000)	4.0			
Model 5	$s_{10}(\ln x, W \ln x)$	4.5 (0.000)	14.7	83.3	-1,606	0.013
	$s_{11}\left(\ln\left[\frac{i_k}{n+g+\delta}\right], W \ln\left[\frac{i_k}{n+g+\delta}\right]\right)$	1.8 (0.017)	10.2			
	$s_{12}(\ln agri, W \ln agri)$	4.5 (0.012)	2.0			
	$s_{13}(\ln dens, W \ln dens)$	3.1 (0.000)	20.1			
	$s_{14}(W\gamma)$	9.5 (0.000)	1.0			

Notes:

- **Model 1:** Linear Solow model; **Model 2:** Nonlinear Solow model; **Model 3:** Spatial nonlinear Solow model; **Model 4:** Nonlinear Solow model augmented with sectoral specialization and employment density; **Model 5** Full model.
- **Coefficients** refer to parametric terms.
- **F tests** are used to investigate the overall ('approximate') significance of smooth terms.
- **E.d.f.** (effective degrees of freedom) reflect the flexibility of the model. An e.d.f. equals to 1 suggests that the smooth term can be approximated by a linear term.
- **Deviance** is the proportion of deviance explained.
- **AIC** is the ('approximate') Akaike Information Criterion.
- The **GCV score** (x 1000) provides a criterion for choosing the model specification among several different possible alternatives. Thus, the decision to remove or maintain a term is based on comparison of GCV scores and the model which minimizes the GCV is preferred.

Table 2 - Diagnostics

		Model 1	Model 2	Model 3 (IV)	Model 4	Model 5 (IV)
Normality	JB (p-values)	33.719 (0.000)	41.408 (0.000)	22.299 (0.000)	11.000 (0.004)	0.063 (0.968)
Spatial depend.	p-values	424 km: 0.000 924 km: 0.000	424 km: 0.000 924 km: 0.000	424 km: 0.253 924 km: 0.481	424 km: 0.000 924 km: 0.000	424 km: 0.313 924 km: 0.948
Constant variance	F test (p-values)	0.725 (0.365)	2.161 (0.026)	1.569 (0.137)	1.467 (0.163)	1.548 (0.144)
F-test on instruments	F test (p-values)			4.972 (0.000)		6.980 (0.000)
Sargan test	Chi-square (p-values)			27.1 (0.076)		42.0 (0.225)

Notes:

- **Model 1:** Linear Solow model; **Model 2:** Nonlinear Solow model; **Model 3:** Spatial nonlinear Solow model; **Model 4:** Nonlinear Solow model augmented with sectoral specialization and employment density; **Model 5** Full model.
- The **normality test** is based on Jarque-Bera (JB) statistics.
- The tests of **spatial dependence** (using two different distance neighbors weights matrices) are based on a Monte Carlo Simulation of Moran's I.
- The **test of constant variance** of the residuals is based on the estimation of the simple model $|\hat{\epsilon}| = \alpha + s(\hat{y}) + \epsilon$, where $|\hat{\epsilon}|$ is the absolute value of the residuals of the model and \hat{y} is the vector of fitted values. Under the null hypothesis of constant variance, the smooth term $s(\hat{y})$ must be estimated with one degree of freedom and, according to a F test, should not have a significant effect on $|\hat{\epsilon}|$.
- The **F-tests on instruments** reports the significance tests for the instruments in the first stage.
- The **Sargan** test is a test of the joint null hypothesis that the excluded instruments are valid instruments, that is a test of the null hypothesis of zero correlation of instruments with second stage residuals (overidentifying restriction) (see Sargan, 1958). The Sargan statistics is easily computed as $u'Hu / (u'u / N)$, where u are the residuals from the IV estimation, H is the "projection matrix" $Z(Z'Z)^{-1}Z'$, with Z the full set of instruments and N is the number of observations.

Figure 1 – Univariate density and ergodic distribution
Local likelihood density estimation with variable bandwidth (Loader, 1996)

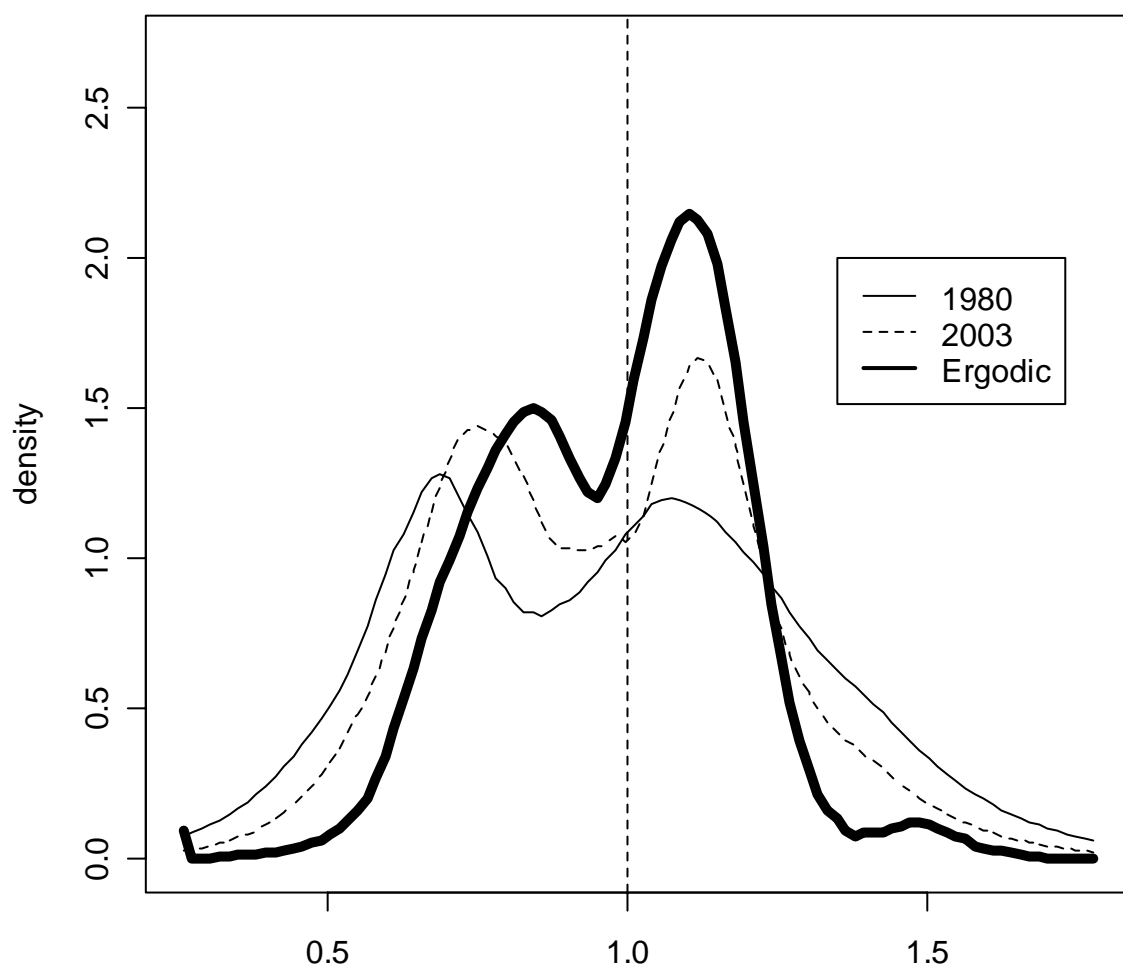
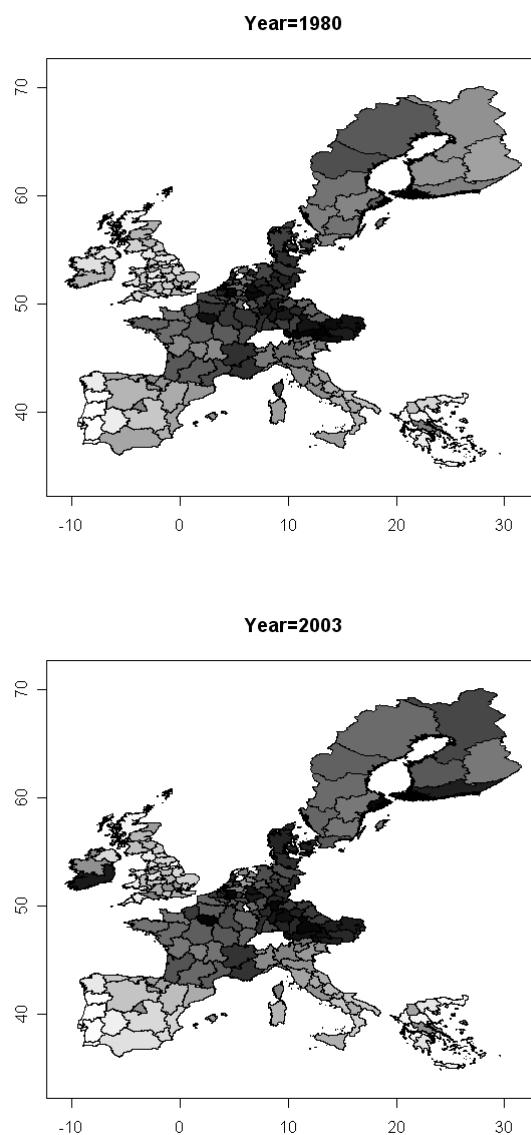


Figure 2 – Choropleth maps of the distribution of labour productivity in 1980 and 2003



Note: regional productivity levels have been classified using 100 break-points. The intensity of the grey color varies with the variable of interest.

Figure 3 – Global G statistics

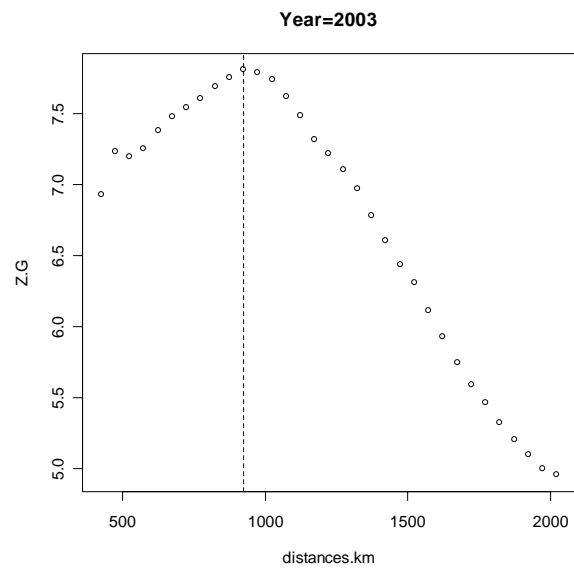
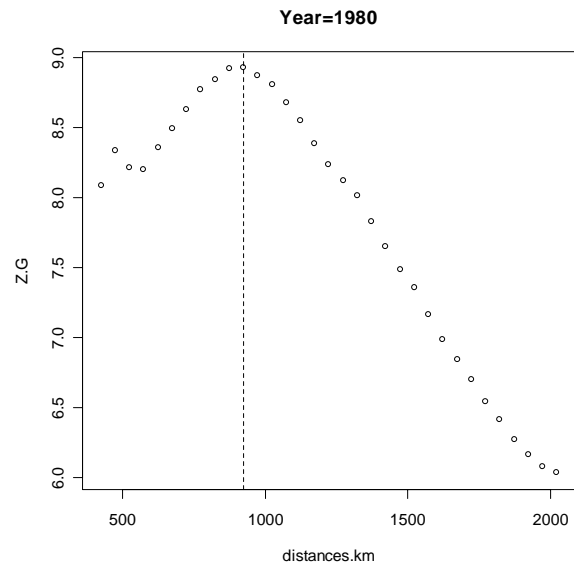


Figure 4 – Maps of G^* indices

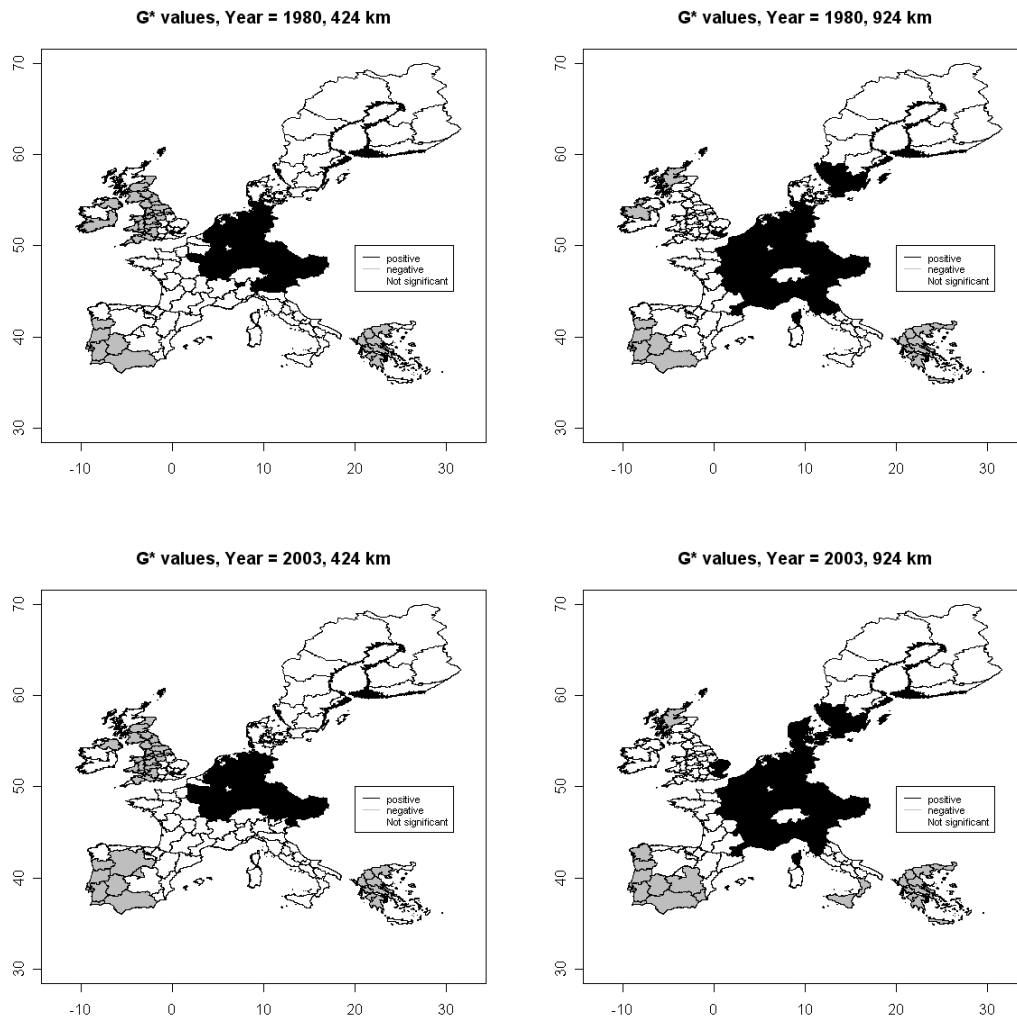
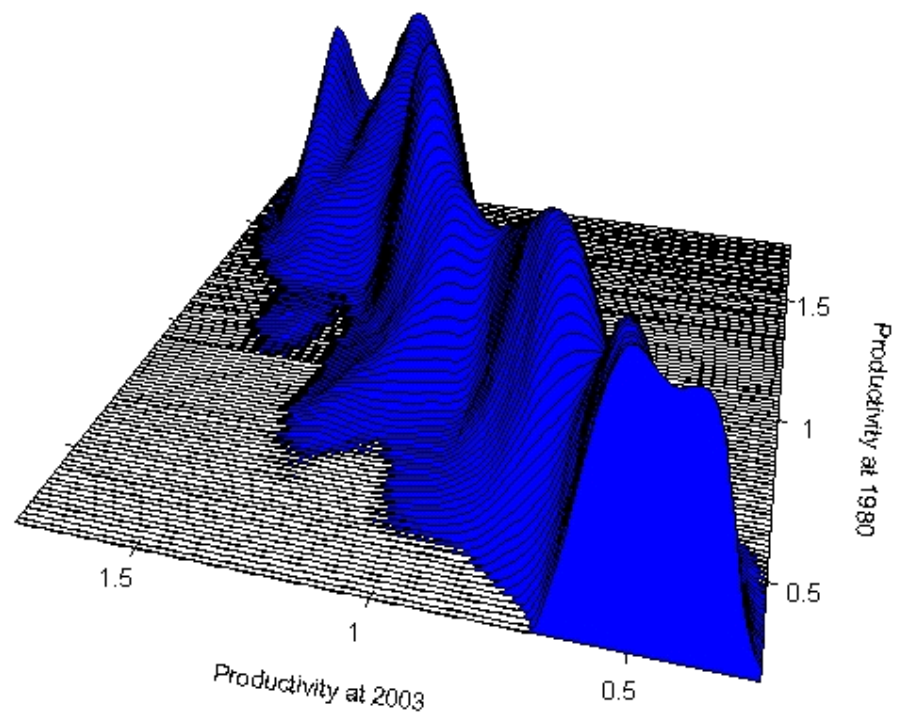
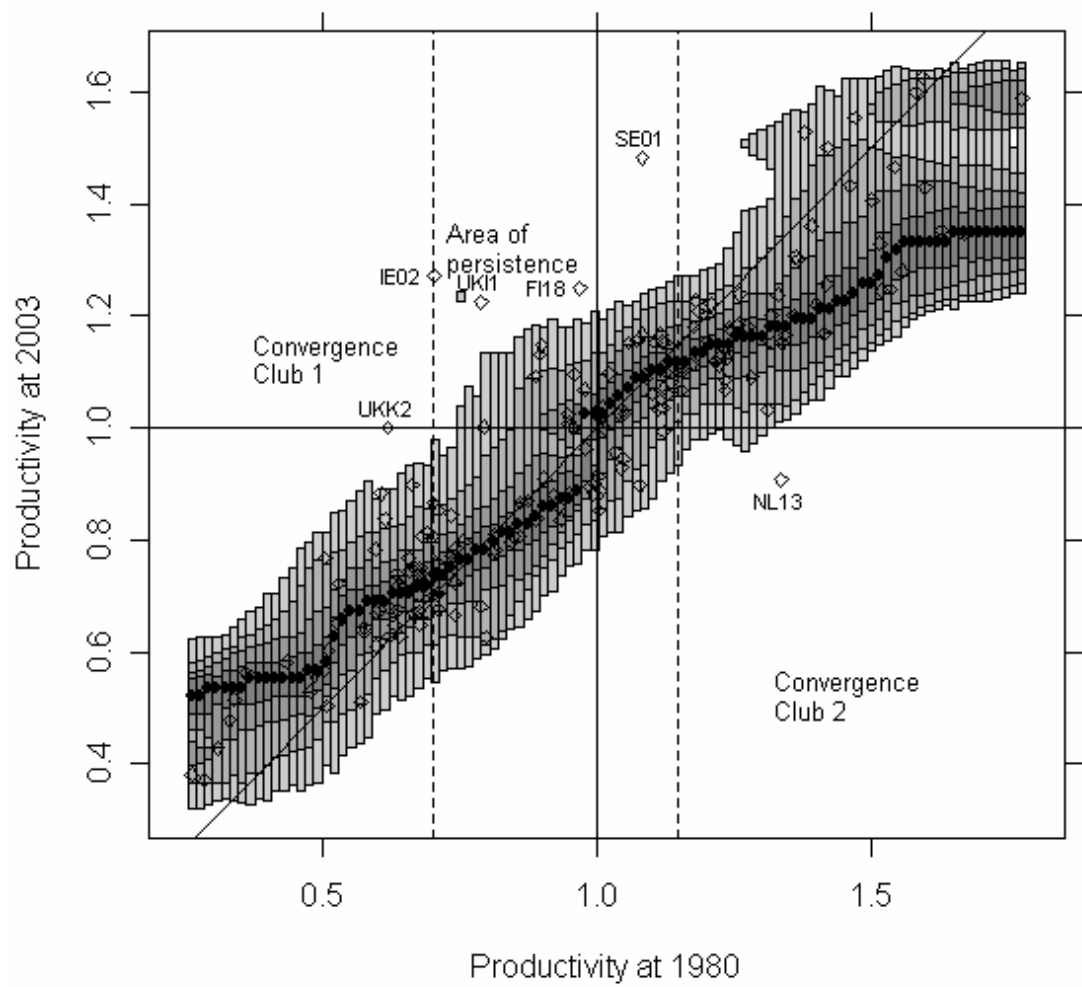


Figure 5 - Intra-Distribution Dynamics

Stacked density plot and HDR plot of conditional density based on the local parametric estimator with variable bandwidth (Hyndman and Yao, 2002)



A) Stacked density plot



B) HDR plot

Figure 6 – The two clubs of convergence

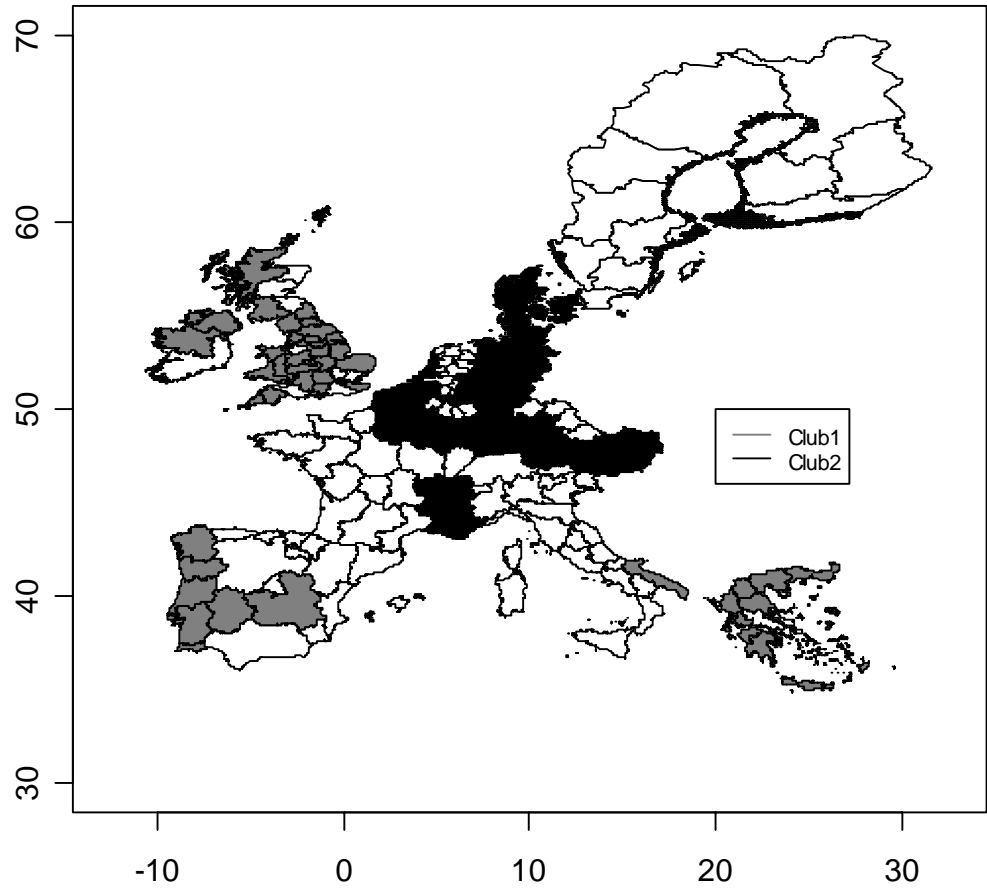
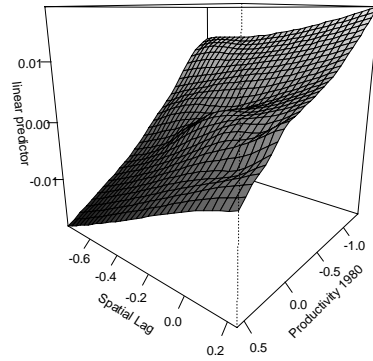
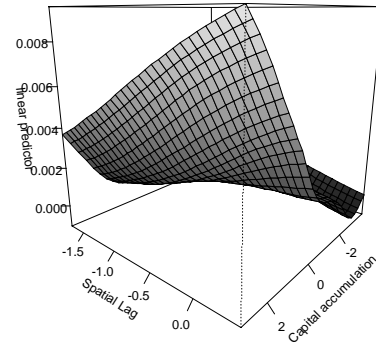


Figure 7 – Growth determinants (model 5)

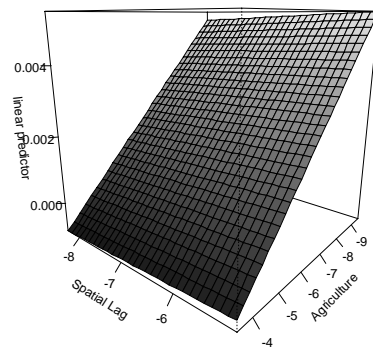
A) Initial conditions



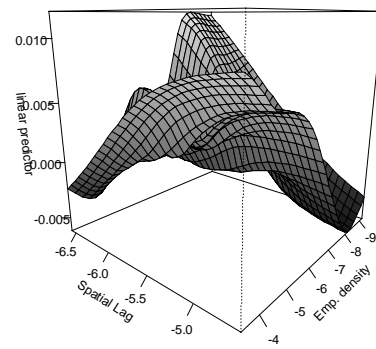
B) Physical capital accumulation



C) Agriculture



D) Employment density



E) Spatial externalities

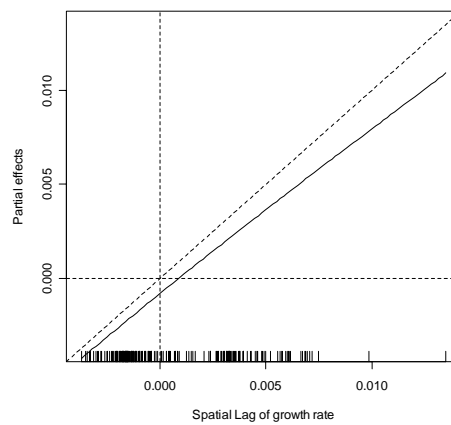
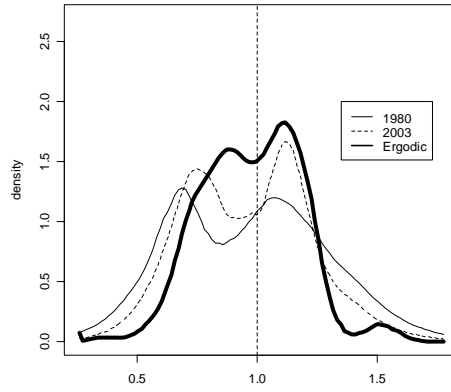
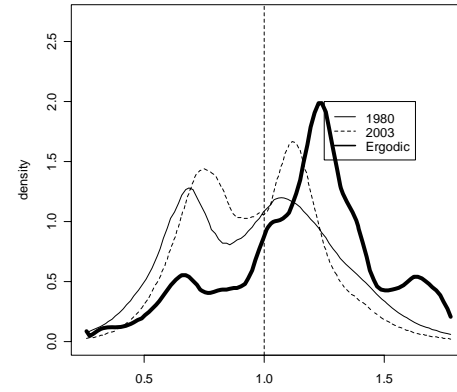


Figure 8 – Ergodic distribution conditioning on growth determinants

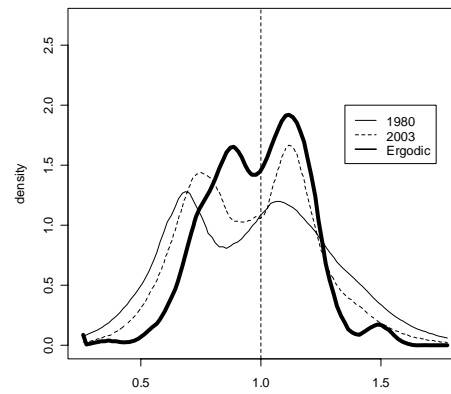
A) Conditioned on model 2



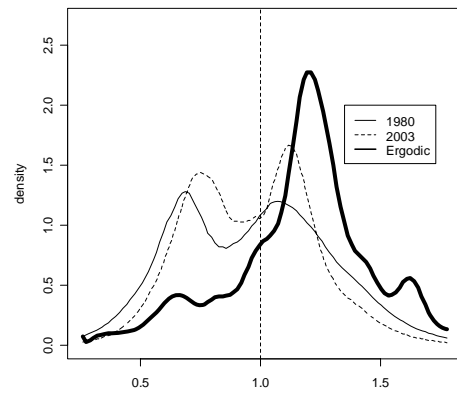
B) Conditioned on model 3 (cutoff: 924 km)



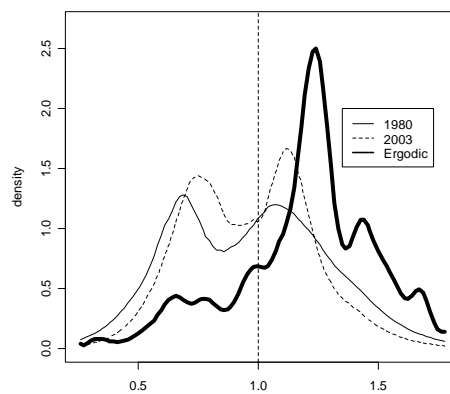
C) Conditioned on model 4



D) Conditioned on model 5 (cutoff: 924 km)



E) Conditioned on model 5 (cutoff: 424 km)



Bibliography

- Anselin L. 1988. *Spatial Econometrics*. Dordrecht: Kluwer Academic Publishers.
- Anselin L. 2004. *Spatial Externalities, Spatial Multipliers and Spatial Econometrics*. *International Regional Science Review* 26, 153-166.
- Armstrong H.W. 1995. An Appraisal of the Evidence from Cross-Sectional Analysis of the Regional Growth Process within the European Union. In: Vickerman RW, Armstrong H.W. (eds.) *Convergence and Divergence Among European Regions*. Pion Limited, London: 40-65.
- Azariadis C., Drazen A. 1990. Threshold Externalities in Economic Development. *Quarterly Journal of Economics* 105, 501-526.
- Azariadis C., Stachurski J. 2004. Poverty Traps. In *Handbook of Economic Growth*. Aghion P. and Durlauf S. (eds.), Amsterdam: North-Holland.
- Basile R. 2007a. Intra-distribution dynamics of regional per-capita income in Europe: evidence from alternative conditional density estimators, *ISAE wp*, 75.
- Basile R. 2007b. Regional Economic Growth in Europe: a Semiparametric Spatial Dependence Approach. *Papers in Regional Science*. forthcoming.
- Bandyopadhyay S. 2003. Polarisation, Stratification and Convergence Clubs: Some Dynamics and Explanations of Unequal Economic Growth across Indian states, mimeo.
- Banerjee A. Duflo E. 2003. Inequality and growth: what can the data say? *Journal of Economic Growth* 8, 267-300.
- Bashtannyk D., Hyndman R. 2001. Bandwidth Selection for Kernel Conditional Density Estimation. *Computational Statistics and Data Analysis* 36, 279-98.
- Bowles S., Durlauf S.N. and Hoff K. 2004. *Poverty Traps*, Princeton University Press.

- Brock W., Durlauf S. 2000. Growth economics and reality. *Working paper*, Department of Economics, University of Wisconsin.
- Ciccone A., Hall R.E. 1996. Productivity and the Density of Economic Activity. *American Economic Review* 86, 54-70.
- Cheshire P., Magrini S. 2006. European Urban Growth: now for some problems of spaceless and weightless econometrics. *Working Papers* 23_06, University of Venice "Ca' Foscari", Department of Economics
- Durlauf S., Johnson P. 1995. Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10, 365-384.
- Durlauf S., Kourtellos A., Minkin A. 2001. The local Solow growth model. *European Economic Review* 45, 928-940
- Durlauf S.N., Johnson P.A., Temple J.R.W. 2005. Growth Econometrics. In *Handbook of Economic Growth*, Volume 1A, Aghion P, Durlauf SN (eds.), North-Holland: Amsterdam.
- Efron B., Tibshirani R. 1993. *An introduction to the bootstrap*. London: Chapman and Hall.
- Ertur C., Koch W. 2006. Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence, *Journal of Applied Econometrics*, forthcoming.
- Fan J, Yao Q, Tong H. 1996. Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika* 83, 189–206.
- Fernàndez C., Ley E., Steel M.F.J. 2002. Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics* 16, 563-576.
- Fiaschi D., Lavezzi M. 2007. Productivity Polarization and Sectoral Dynamics in European Regions, *mimeo*.

- Fotheringham, A.S., Brunson, C., Charlton M. 2000. *Quantitative Geography: Perspectives on Spatial Data*, Sage Publications Ltd.
- Galor O. 1996. Convergence? Inferences from theoretical models. *Economic Journal* 106, 1056-1069
- Getis, A, Ord, J.K. 1992. The analysis of spatial association by use of distance statistics, *Geographical Analysis* 24, 195-.
- Hall P., Wolff R., Yao Q. 1999. Methods for estimating a conditional distribution function. *Journal of American Statistical Association* 94, 154–163.
- Hastie T.J., Tibshirani R.J. 1990. *Generalized Additive Models*. New York: Chapman and Hall
- Hyndman R.J. 1996. Computing and Graphing Highest Density Regions. *The American Statistician* 50, 120-126.
- Hyndman R.J., Bashtannyk D.M., Grunwald G.K. 1996. Estimating and visualizing conditional densities. *Journal of Computational and Graphical Statistics* 5, 315-336.
- Hyndman R.J., Yao Q. 2002. Nonparametric estimation and symmetry tests for conditional density functions. *Journal of Nonparametric Statistics* 14, 259-278
- Kelejian H.H., Prucha I.R. 1998. A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17, 99-121
- Kurtellos A. 2001. Modelling parameter heterogeneity in cross-country growth regression models. *Working paper*, Department of Economics, University of Wisconsin.
- Lamo A. 2000. On Convergence Empirics: Some Evidence for Spanish Regions. *Investigaciones Economicas* 24, 681-707.

- Leonida L. 2003. On The Effects of Industrialization on Growth and Convergence Dynamics in Italy (1960-95), *mimeo*.
- Liu Z., Stengos T. 1999. Non-Linearities in Cross-Country Growth Regressions: a Semiparametric Approach. *Journal of Applied Econometrics* 14, 527-538.
- Loader C.R. 1996. Local likelihood density estimation. *The Annals of Statistics* 24, 1602–1618.
- López-Bazo E., Vayà E., Artis M. 2004. Regional externalities and growth: evidence from European regions. *Journal of Regional Science* 44, 43-73.
- Magrini S. 2004. Regional (Di)Convergence. In *Handbook of Regional and Urban Economics*, Henderson V, Thisse JF (eds.). North-Holland: Amsterdam.
- Masanjala W.H., Papageorgiou C. 2004. The Solow model with CES technology: nonlinearities and parameter heterogeneity. *Journal of Applied Econometrics* 19, 171-201.
- Ord J.K., Getis A. 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis* 27, 286–306.
- Paci R. 1997. More similar and less equal. Economic growth in the European regions, *Weltwirtschaftliches Archiv*, 4.
- Paci R., Pigliaru F. 1999. European Regional Growth: Do Sectors Matter? In Adams, J. and F. Pigliaru (ed.) *Economic Growth and Change. National and Regional Patterns of Convergence and Divergence* Cheltenham, UK: Edward Elgar.
- Pittau M.G., Zelli R. 2006. Income dynamics across EU regions: empirical evidence from kernel estimator. *Journal of Applied Econometrics* 21, 605-628.
- Quah D. 1993. Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics* 95, 427–443.

- Quah D. 1996a. Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *Economic Journal* 106, 1045-55.
- Quah D. 1996b. Empirics for economic growth and convergence. *European Economic Review* 40, 1353–1375.
- Quah D. 1996c. Convergence Emprics Across Economies with (Some) capital Mobility. *Journal of Economic Growth* 1, 95-124.
- Quah D. 1997. Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth* 2, 27-59.
- Quah D. 2006. Growth and distribution. *Mimeo LSE Economc department* August.
- Rodríguez-Pose A. 1999. Convergence or divergence? Types of Regional Responses to Socioeconomic Change. *Journal of Economic and Social Geography* 90: 363-378.
- Sampson R.J., Morenoff J.D. 2004. Durable inequality: spatial dynamics, social processes, and the persistence of poverty in Chicago neighborhoods. In Bowles S., Durlauf S.N. and Hoff K. 2004. *Poverty Traps*, Princeton University Press (Chap. 7).
- Sargan J.D. 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26: 393--415
- Johnson PA. 2004. *A continuous state space approach to “convergence by parts”*. Department of Economics, Vassar College, Poughkeepsie, NY.
- Wood S.N. 2000. Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties. *Journal of Royal Statistical Society Series B* 62(2): 413-428
- Wood S.N. 2003. Thin plate regression splines. *Journal of Royal Statistical Society Series B* 65(1):95-114.
- Wood S.N. 2006. *Generalized Additive Models. An Introduction with R*. Boca Ratom: Chapman & Hall/CRC.

Wood S.N., Augustin 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling* 157, 157-177

QUADERNI DEL DIPARTIMENTO DI ECONOMIA, FINANZA E STATISTICA

Università degli Studi di Perugia

1	Gennaio 2005	Giuseppe CALZONI Valentina BACCHETTINI	Il concetto di competitività tra approccio classico e teorie evolutive. Caratteristiche e aspetti della sua determinazione
2	Marzo 2005	Fabrizio LUCIANI Marilena MIRONIUC	Ambiental policies in Romania. Tendencies and perspectives
3	Aprile 2005	Mirella DAMIANI	Costi di agenzia e diritti di proprietà: una premessa al problema del governo societario
4	Aprile 2005	Mirella DAMIANI	Proprietà, accesso e controllo: nuovi sviluppi nella teoria dell'impresa ed implicazioni di corporate governance
5	Aprile 2005	Marcello SIGNORELLI	Employment and policies in Europe: a regional perspective
6	Maggio 2005	Cristiano PERUGINI Paolo POLINORI Marcello SIGNORELLI	An empirical analysis of employment and growth dynamics in the italian and polish regions
7	Maggio 2005	Cristiano PERUGINI Marcello SIGNORELLI	Employment differences, convergences and similarities in italian provinces
8	Maggio 2005	Marcello SIGNORELLI	Growth and employment: comparative performance, convergences and co-movements
9	Maggio 2005	Flavio ANGELINI Stefano HERZEL	Implied volatilities of caps: a gaussian approach
10	Giugno 2005	Slawomir BUKOWSKI	EMU – Fiscal challenges: conclusions for the new EU members
11	Giugno 2005	Luca PIERONI Matteo RICCIARELLI	Modelling dynamic storage function in commodity markets: theory and evidence
12	Giugno 2005	Luca PIERONI Fabrizio POMPEI	Innovations and labour market institutions: an empirical analysis of the Italian case in the middle 90's
13	Giugno 2005	David ARISTEI Luca PIERONI	Estimating the role of government expenditure in long-run consumption
14	Giugno 2005	Luca PIERONI Fabrizio POMPEI	Investimenti diretti esteri e innovazione in Umbria
15	Giugno 2005	Carlo Andrea BOLLINO Paolo POLINORI	Il valore aggiunto su scala comunale: la Regione Umbria 2001-2003

16	Giugno 2005	Carlo Andrea BOLLINO Paolo POLINORI	Gli incentivi agli investimenti: un'analisi dell'efficienza industriale su scala geografica regionale e sub regionale
17	Giugno 2005	Antonella FINIZIA Riccardo MAGNANI Federico PERALI Paolo POLINORI Cristina SALVIONI	Construction and simulation of the general economic equilibrium model Meg-Ismea for the italian economy
18	Agosto 2005	Elżbieta KOMOSA	Problems of financing small and medium-sized enterprises. Selected methods of financing innovative ventures
19	Settembre 2005	Barbara MROCZKOWSKA	Regional policy of supporting small and medium-sized businesses
20	Ottobre 2005	Luca SCRUCCA	Clustering multivariate spatial data based on local measures of spatial autocorrelation
21	Febbraio 2006	Marco BOCCACCIO	Crisi del welfare e nuove proposte: il caso dell'unconditional basic income
22	Settembre 2006	Mirko ABBRITTI Andrea BOITANI Mirella DAMIANI	Unemployment, inflation and monetary policy in a dynamic New Keynesian model with hiring costs
23	Settembre 2006	Luca SCRUCCA	Subset selection in dimension reduction methods
24	Ottobre 2006	Sławomir I. BUKOWSKI	The Maastricht convergence criteria and economic growth in the EMU
25	Ottobre 2006	Jan L. BEDNARCZYK	The concept of neutral inflation and its application to the EU economic growth analyses
26	Dicembre 2006	Fabrizio LUCIANI	Sinossi dell'approccio teorico alle problematiche ambientali in campo agricolo e naturalistico; il progetto di ricerca nazionale F.I.S.R. – M.I.C.E.N.A.
27	Dicembre 2006	Elvira LUSSANA	Mediterraneo: una storia incompleta
28	Marzo 2007	Luca PIERONI Fabrizio POMPEI	Evaluating innovation and labour market relationships: the case of Italy
29	Marzo 2007	David ARISTEI Luca PIERONI	A double-hurdle approach to modelling tobacco consumption in Italy
30	Aprile 2007	David ARISTEI Federico PERALI Luca PIERONI	Cohort, age and time effects in alcohol consumption by Italian households: a double-hurdle approach
28	Marzo 2007	Luca PIERONI Fabrizio POMPEI	Evaluating innovation and labour market relationships: the case of Italy

29	Marzo 2007	David ARISTEI Luca PIERONI	A double-hurdle approach to modelling tobacco consumption in Italy
30	Aprile 2007	David ARISTEI Federico PERALI Luca PIERONI	Cohort, age and time effects in alcohol consumption by Italian households: a double-hurdle approach
31	Luglio 2007	Roberto BASILE	Productivity polarization across regions in Europe

I QUADERNI DEL DIPARTIMENTO DI ECONOMIA
Università degli Studi di Perugia

1	Dicembre 2002	Luca PIERONI:	Further evidence of dynamic demand systems in three european countries
2	Dicembre 2002	Luca PIERONI Paolo POLINORI:	Il valore economico del paesaggio: un'indagine microeconomica
3	Dicembre 2002	Luca PIERONI Paolo POLINORI:	A note on internal rate of return
4	Marzo 2004	Sara BIAGINI:	A new class of strategies and application to utility maximization for unbounded processes
5	Aprile 2004	Cristiano PERUGINI:	La dipendenza dell'agricoltura italiana dal sostegno pubblico: un'analisi a livello regionale
6	Maggio 2004	Mirella DAMIANI:	Nuova macroeconomia keynesiana e quasi razionalità
7	Maggio 2004	Mauro VISAGGIO:	Dimensione e persistenza degli aggiustamenti fiscali in presenza di debito pubblico elevato
8	Maggio 2004	Mauro VISAGGIO:	Does the growth stability pact provide an adequate and consistent fiscal rule?
9	Giugno 2004	Elisabetta CROCI ANGELINI Francesco FARINA:	Redistribution and labour market institutions in OECD countries
10	Giugno 2004	Marco BOCCACCIO:	Tra regolamentazione settoriale e antitrust: il caso delle telecomunicazioni
11	Giugno 2004	Cristiano PERUGINI Marcello SIGNORELLI:	Labour market performance in central european countries
12	Luglio 2004	Cristiano PERUGINI Marcello SIGNORELLI:	Labour market structure in the italian provinces: a cluster analysis
13	Luglio 2004	Cristiano PERUGINI Marcello SIGNORELLI:	I flussi in entrata nei mercati del lavoro umbri: un'analisi di cluster
14	Ottobre 2004	Cristiano PERUGINI:	Una valutazione a livello microeconomico del sostegno pubblico di breve periodo all'agricoltura. Il caso dell'Umbria attraverso i dati RICA-INEA
15	Novembre 2004	Gaetano MARTINO Cristiano PERUGINI	Economic inequality and rural systems: empirical evidence and interpretative attempts
16	Dicembre 2004	Federico PERALI Paolo POLINORI Cristina SALVIONI Nicola TOMMASI Marcella VERONESI	Bilancio ambientale delle imprese agricole italiane: stima dell'inquinamento effettivo